Turp, M., González, J.C., Pulido, J.C., Fernández, Fernando. (2019). Developing a robot-guided interactive simon game for physical and cognitive training. *International Journal of Humanoid Robotics*, 16(No. 01)

DOI:https://doi.org/10.1142/S0219843619500038

International Journal of Humanoid Robotics


# Developing a Robot-Guided Interactive Simon Game for Physical and Cognitive Training

Mısra Turp*, José Carlos González[†], José Carlos Pulido[‡] and Fernando Fernández[§]

*Computer Science Department, Universidad Carlos III de Madrid,*
*Av. de la Universidad, 30, 28911 Leganés, Madrid, Spain*
**misraturp@gmail.com*
*[†]josgonza@inf.uc3m.es*
*[‡]jcpulido@inf.uc3m.es*
*[§]ffernand@inf.uc3m.es*

Received Day Month Year
Revised Day Month Year
Accepted Day Month Year

Enveloping cognitive or physical rehabilitation into a game highly increases the patients' commitment with their treatment. Specially with children, maintaining them motivated is a very time-consuming work, so therapists are demanding tools to help them with this task. NAOTherapist is a generic robotic architecture that uses Automated Planning techniques to autonomously drive non-contact upper-limb rehabilitation sessions for children with a humanoid NAO robot. Our aim is to develop more robotic games for this platform to enrich its variability and possibilities of interaction. The goal of this work is to present our first attempt to develop a different, more complex game that reuses the previous architecture. We contribute with the design description of a novel robotic Simon game that employs upper-limb poses instead of colors and could qualify as a cognitive and physical training. Statistics of evaluation tests with 14 adults and 56 children are displayed and the outcomes are analyzed in terms of human-robot interaction quality. The results demonstrate the application-domain generalization capabilities of the NAOTherapist architecture and give an insight to further analyze the therapeutic benefits of the new developed Simon game.

*Keywords*: Human-Robot Interaction; Automated Planning; Interactive Games; Cognitive Robotics; Robotic Therapy

## 1. Introduction

In general, playing is basic for children's development to gather a lot of physical and social stimuli. Some cognitive disorders can limit their playing time, with or without other children, which is problematic specially at certain key ages.[1] Rehabilitation and cognitive therapies can be boring and repetitive, and sometimes they must be performed in a small time-window during childhood. This can cause disengagement of the patients, reducing the effectiveness of their treatment. Specially with children, transforming a therapy into a game (also called gamification or serious games) has become an important way to maintain them motivated.[2,3]

Socially Assistive Robotics (SAR) is a very active research field that is bringing actual social robots to environments like rehabilitation and cognitive therapies.[4] It focuses on robots to assist people through social interaction rather than physical interaction. There are several open challenges and limitations in the physical world that currently prevent their massive use and commercialization, in contrast with videogames or virtual avatars. The impact of the interaction with a social physical entity like a robot is very intense, specially if it has a good level of embodiment.[5] Sometimes, the mere presence of the robot can improve even the engagement of children with autism.[6,7,8] Such intensity is more difficult to achieve with videogames due to their virtual nature.[9] Anyway, motivating and eye-catching elements like virtual avatars and social robots are useful to encourage interactions with these children and stimulate them.[10]

In this context, the NAOTherapist architecture employs a humanoid NAO robot to autonomously drive upper-limb rehabilitation sessions for children.[11] It implements the mirror game, which is an imitation activity where the child has to mimic the posture of the robot. The child is detected through a 3D sensor and the robot indicates him how to correct his posture, if needed. This robot is intended to be used in hospitals, clinics, rehabilitation centers, etc., because it also gathers clinical information constantly that can be used by a therapist to assess the evolution of the patient. NAOTherapist has been already evaluated with healthy children and three pediatric patients,[12] and the outcome was very promising. However, in order to engage children much more in long-term treatments, it is needed to increase the interactive possibilities of the robotic platform.

Our long-term goal is to create a diverse catalogue of games for the NAOTherapist platform. To do that, the reuse of previous components is essential, since the architecture is application independent. This work describes exhaustively and expands our first attempt to design a new game for the existing NAOTherapist architecture: an interactive game application based on the mechanism of the popular electronic memory skill game Simon by using a NAO robot as a coach.[13] To the best of our knowledge, this is the first time that the Simon game has been developed with poses instead of colors and with a humanoid robot performing them, all controlled by a cognitive robotic architecture.

The new game requires not only physical abilities to perform the poses, but also memory skills to remember the pose sequences. This cognitive addition increases the potential users of NAOTherapist, because initially it was intended only for patients with upper-limb disorders. Reusing the same robotic architecture allows a much faster and easier development, even to the point that new activities could jump from the clinical environment to just mere entertainment on a daily basis or even be adapted to an adult or elder profile.

An important aspect of this work is that NAOTherapist uses Automated Planning as the core for its deliberative processes.[14] This artificial intelligence technique uses only descriptions of the available actions, the current state of the world and some goals to generate a suitable plan of actions. These descriptions are declarative models written in Planning Domain Definition Language (PDDL),[15] which eliminate the need to maintain a big and coherent finite-state machine. Thus, to create a different application domain like a new game in this work, it is needed to model it in PDDL.

This work describes the formalized model of the Simon game based on Automated Planning. This model is technically more complex than the original mirror game, as it contains two execution loops and a shuffling mechanism to create unpredictable pose chains, which are required for Simon. Two different evaluations were carried out to assess the Human-Robot Interaction (HRI) quality of the developed game. The first one was done with 14 adults to discover interaction issues with the system and to make it more suitable for children. In the second evaluation, the game was tested with 56 children of ages 5 and 6 to assess the performance of the game itself and the interaction quality of the robot. The outcome of both tests successfully demonstrates the domain generalization capabilities of NAOTherapist and provides a positive outlook for the future use of the Simon game with therapeutic objectives, which are out of the scope of this manuscript.

Section 2 describes related work about robots used in rehabilitation and robotic control architectures. Section 3 explains technical aspects of the original NAOTherapist architecture. Sections 4 and 5 describe the robotic Simon use case and their implementation within the architecture. Section 6 details the HRI evaluations with adults and children, while in sections 7 and 8 the results are discussed and several conclusions are presented addressing the current and future state of the project.

## 2. Related Work

Aside from social robots for rehabilitation, virtual avatars and videogames are much easier to develop and deploy.[16,17] They lack the impact of a physical presence,[5] but still, they can be helpful specially for cognitive rehabilitation. However, some physical activities can be better managed with an actual robot. For instance, using a robot for sign-language activities allows the user to check the hand pose at every angle,[18] which is the most natural way to train them. Arm poses of the developed Simon game can benefit from this important aspect too.

Autonomy is a major issue while trying to apply these solutions in daily life. Robots can be a useful tool as a playmate even if they are teleoperated.[19] However, their lack of autonomy is a severe drawback. More autonomous approaches like Leonardo focus on joining social interaction and gaming to increase the children's commitment with the therapy.[20] This robot manages to do it autonomously in spite of lacking voice. A truly autonomous interaction should not rely on any wearable device, although there are approaches that use this kind of interaction, like gloves.[21]

Automated Planning techniques can be used to support autonomy. Among them, Hierarchical Task Networks (HTN) exploit the hierarchical nature of many robotic behaviors.[22] Some works use them for the robot control or for the high-level task specification such as rehabilitation plans,[23,24] but HTNs are not as domain independent as the approach presented here.

Social robotics sometimes is specialized in a particular application domain, like physical rehabilitation.[25] Some of them are designed for children with very specific problems like brachial plexus palsy or cerebral palsy.[26,27,28] On the contrary, NAOTherapist was designed to be more domain independent.[11] This manuscript demonstrates its generalization capabilities that allow it to work for cognitive or physical training.

## 3. NAOTherapist architecture

This work is an extension of the NAOTherapist architecture.[11] An architecture for Human-Robot Interaction (HRI) is very important for the success of an autonomous robotic platform because there are many different challenges that must be solved like artificial vision, situation awareness, behavior control, etc. These architectures are the essential structure of a domain-generic computational cognitive model.[29] Thus, NAOTherapist was designed as a modular architecture. It uses the RoboComp framework to organize its components,[30] similar to ROS.[31] They allow a much simpler integration of all these components and eases their replacement, extension or reuse for other application domains.

A diagram of the general structure of the architecture can be found in Section 5.2, but as a summary it is as follows: an Executive component joins Vision, Robot and Decision Support components. Vision is connected to a Kinect 3D sensor and Robot to a NAO robot. Decision Support is in charge of controlling the behavior of the robot depending on the current state of the detected environment. Execution takes information from Vision and Robot to create an updated state (user is in front of the sensor and the robot is standing, for instance) and calls Decision Support for new "actions" to execute when the robot is idle.

Decision Support integrates an Automated Planning and monitoring subarchitecture called PELEA.[32] In Automated Planning,[14] an automated planner uses advanced heuristics to search for a plan of actions that transforms an initial state into a final state that accomplishes certain goals. Each action checks the current state to determine if it can be executed or not and specifies the predicted effects over the state. All actions are defined in a domain file written in PDDL.[15] The initial state and the goals are defined in the problem file, also in PDDL language.

PELEA is useful to replan accordingly to unexpected situations (user leaves the training area or the pose is wrong, for instance). Decision Support receives the updated state and evaluates if it is compatible with the previously generated plan of actions. If so, the component sends the next action to Execution. If it is not compatible, it calls to the automated planner again and replans, sending the first action of the new plan to Execution.

The actions managed by the automated planner must be decomposed later into instructions for the robot to execute them, such as to wave to greet the user, move arms to show the pose, say something, etc. This is performed in the Executive component.

Finally, regarding to the pose detection, the Vision component is in charge of managing the 3D sensor. It checks how similar the actual user's pose is with respect to the expected one. Expected poses are stored into a catalogue in a knowledge base. The comparison uses the angles of the joints as a distance metric and updates a dynamic acceptance threshold according to the performance of the user to avoid frustration.[12]

So, to create a new game like Simon, the components of the architecture that must be updated are two: the Decision Support to model all the required behavior in PDDL; and the Executive component to model the decomposition of each action into a set of low-level instructions.

On a side note, this manuscript ignores the high level of planning of the original architecture, which comprises a high-level therapy designer to generate a clinically suitable schedule of sessions with exercises (high-level actions) for a patient.[11] The Simon game is focused on the poses of these exercises, the interaction with the patient and replanning (medium-level actions). Thus, for clarity purposes, in this manuscript the term "actions" refers to the concept of medium-level actions, while "instructions" refers to low-level robot actuators.

## 4. Robotic Simon With Poses

The original Simon is a game played with an interactive electronic device whose purpose is to train the short-term memory. During this game, a sequence with an increasing number of colors is displayed on the device and the player is expected to repeat the same pattern. The game goes on until the player makes a mistake in the order of the colors. It has been also used in psychology research as a measure of working memory span and as a tool for training and predicting memory performance.[33,34] Not to be confused with the Simon-says game, in which a player gives commands preceded by the phrase "Simon says" that the other players must perform quickly.[26] This latter game trains reflexes instead of memory skills and it is not covered in this manuscript.

### 4.1. *Arm poses instead of colors*

In this work, we propose a new point of view of the traditional electronic Simon game by merging both physical and cognitive training driven by a humanoid robot. To do this, the game flow changes by using upper-arm body poses instead of colors and the game lasts for a pre-determined number of turns. In contrast to the original game, a humanoid robot reminds the pose to the user once for ease of play (in case he forgets the sequence or how exactly a certain pose was), since it is much more complicated to remember arm positions than colors. The chances given to the user can be raised or lessened to adapt them to the age, cognitive capability and physical constraints of the player. As an example, Figure 1 shows part of a game session during the success and failure of a user and some vocalizations provided by the NAO robot.

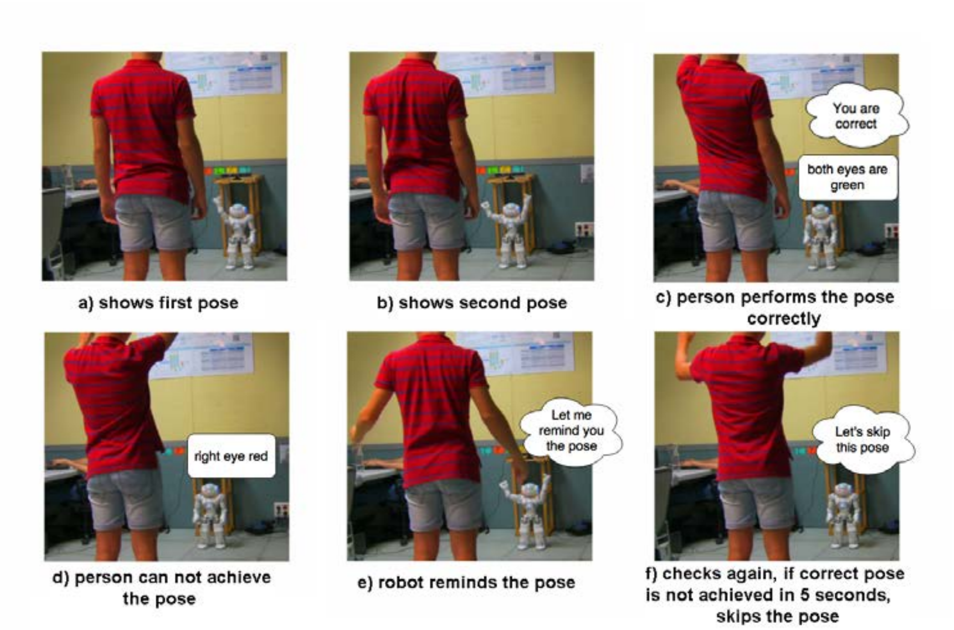



Fig. 1. Course of the robot-guided Simon game helping the user with clues and reminders when the pose is wrong.

The initial goal was to use this project on pediatric patients to show them a game-based rehabilitation process and take them out of the boring routine exercise ambiance. However, on the development process, the system was also tested with adults to evaluate their perception about the speed of the game, the quality of the given feedback and the difficulty of the poses.

Several considerations were taken for the design. It was decided to have most of the poses using one arm at a time or symmetrical ones, which greatly lowers the risk to over-complicate people. But even after lowering the difficulty of the poses, due to the challenge of remembering the sequence, the robot must give the user enough time/tries to achieve the current pose.

While designing the new instructions with the animations and speech of the robot, the primary aim was to keep the communication with the user as simple and clear as possible to avoid confusions. Any feedback that may overwhelm or depress will lower the quality of the interaction dramatically. Thus, encouraging and supportive language is used throughout the session.

In addition to the feedback, the robot should not be acting too slow to avoid the risk of losing interest nor too fast to confuse the player. The design of the game speed tries to emulate a natural conversation to please most of the users, but it can be easily altered if further experimental observations prove the necessity. Overall, the goal is to increase the people's motivation to play with the robot again.

### 4.2. *Use case of the new game*

Figure 2 illustrates the game flow which is arranged briefly as follows: when the user is detected the robot starts the game. The first sequence of actions is executed with only one pose, then the pose is performed by the user and it is checked by the system. If it is correct, the robot moves on to the next sequence that has two consecutive poses (first one being the pose executed on the first sequence). Then it goes on with an increasing number of poses, adding a new pose at the end of the former sequence. Each new added pose is shuffled among the poses declared in the problem file. However if the user does not perform a pose correctly, the robot takes extra actions including reminders and skipping the pose if needed.

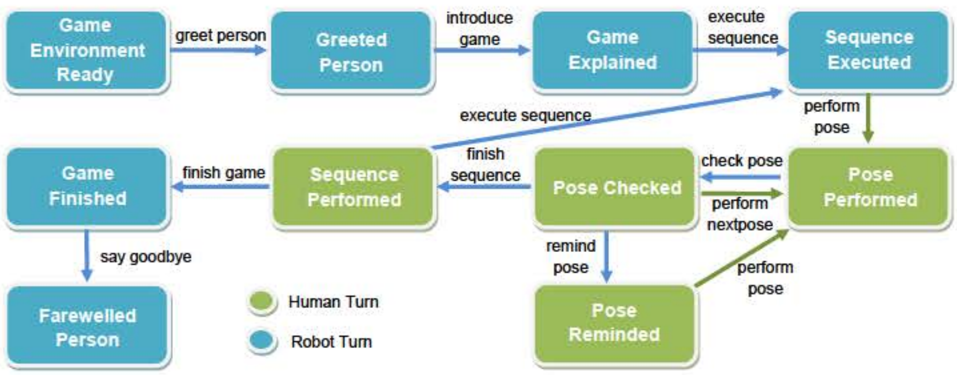



Fig. 2. Simon game execution flow.

For the beginning of the game, the user stands about 2 meters away from the robot and a Microsoft Kinect 3D sensor behind it.

As it was shown in Figure 1, in the checking process where the player is trying to perform the pose, a hint through the eyes of the robot is given, where green indicates the right pose and red a wrong one. The color of the eyes is instantly changed to the tones between red and green in order to give a visual feedback to the user about how close he is to achieve the correct pose. For convenience, both eyes are distinctly colored to indicate the correctness of the different arms. The right eye is for the left arm of the patient and left eye is for the right arm because the poses of the robot are mirrored. Once the pose is successfully performed, the user sees both eyes green and has to hold the pose until the robot gives a distinct audio feedback to communicate that the pose was correctly performed. After this feedback, the robot waits for the next pose or presents the next sequence.

If the user cannot perform the pose correctly after 10 seconds, the robot tells him that it will remind the pose one last time. Then the robot executes only the next expected pose from the user and waits 5 more seconds to see the right posture. When the pose is accomplished after the reminder, the game goes on as usual. If it is still not correct, the robot skips the pose but does not exclude it from the sequence of poses and it will be executing it again in the next round. Three skipped poses cease the game.

Algorithm 1 represents the basic mechanics of the Simon game. It does not take into account the recovery from unexpected situations like the patient leaving the training area or the robot running out of battery. The input variable *storedPoses* is the pose catalogue that can be shuffled to create a new Simon game. This catalogue is stored into the NAOTherapist knowledge base and consists of 7 different pose types per arm.[12] The input variable *maxPoses* is the maximum length of a sequence and *maxSkips* the maximum number of failed turns in a session (3 by default in the present implementation). The local variable *sequence* increases each time that a sequence is completed. The game ends when *sequence* reaches *maxPoses* or when the local variable *skip* reaches *maxSkips*.

In essence, the whole interaction looks like this: the child stands in front of the robot and it starts telling him some basic instructions about the game. The robot shows the first sequence of only one pose with its arms and then waits for the child to mimic it. While the robot is checking the child's pose, it gives him feedback by changing the color gradient of its eyes and reminding the last pose if needed. The child will hear a sound if the pose was correctly done and the robot will move to the next sequence. The second sequence has two poses, the third has three, etc. The game will end when the maximum length of the sequence is reached or there was too many remindings. A video showing a full Simon game between the NAO robot and a children can be found in the web.[a] This video also have subtitles to follow the main vocalizations of the developed system.

## 5. Implementing the Simon Game into the Robotic Architecture

The Simon game has been implemented within the NAOTherapist architecture,[11] which is an execution and monitoring robotic system based on Automated Planning for the development of rehabilitation exercises.[14] This section explains the first attempt to change the original application domain to check the generalization capabilities of the architecture. The Decision Support and the Executive components required modifications to implement the new Simon game, so the following subsections detail the changes needed for each of them.

[a]Video of the use case: *https://youtu.be/picw9sD5VH4*

**Algorithm 1:** Simon pseudocode within the NAOTherapist architecture

```
Input: storedPoses, maxPoses, maxSkips
poses ← ComputeIds(storedPoses, maxPoses)
GreetUser()
IntroduceGame()
skip ← 0
sequence ← 0
while sequence ≤ maxPoses and skip ≤ maxSkips do
    IntroduceSequence()
    for p ← 0 to sequence do
        Perform(poses[p])
    p ← 0
    while p ≤ sequence and skip ≤ maxSkips do
        if ¬Check(poses[p]) then
            RemindPose()
            Perform(poses[p])
            if ¬Check(poses[p]) then
                skip = skip + 1
                SkipPose()
        p ← p + 1
    sequence ← sequence + 1
FinishGame(sequence)
SayGoodbye()
```

### 5.1. *Simon as an Automated Planning task*

In Simon, poses are the basic elements to manage and, like colors in the original game, each sequence should be apparently random. Decision Support contains the PDDL domain that controls the execution flow and also a pose shuffling mechanism. This mechanism is challenging because the automated planner used in NAOTherapist is deterministic and cannot perform any kind of random selection.

Figure 3 is an example of a pose definition in PDDL code. Poses are initially stated in the problem model together with their identification numbers, the specific indicators of the postures of each arm, the game they belong to and their special pose weights that help the shuffling process. In Figure 3, $p0$ indicates that the left arm should be down in rest position and $p4$ indicates that the right arm must be up straight (the pose is mirrored in the robot).

As mentioned before referring to Algorithm 1, the game flow is mainly driven by two loops working one into another. The PDDL representation of the domain comprises an outer loop to execute the turns that contain the sequences and an inner loop to execute the sequences that contain the poses. In each new turn, a new pose is added to the sequence of the last turn, so the number of poses are increasing every turn. Each game (as a product of different PDDL problems) can last until the user fails a sequence a number of times or until the maximum number of turns is executed, specified by a numeric fluent.

```
(pose_game p_id0 g0)
(pose_postures p_id0 p0 p4)
(= (p-position p_id0) 0)
(= (pose_weight p_id0) 2)
```

Fig. 3. Example of the definition of a pose in a PDDL problem file performed by the NAO robot.

In Simon, each new pose added into the sequence must be random, so it is not possible to guess the next pose nor to remember it with a logical order. The pose shuffling system uses an index value into the PDDL problem that serves as a pseudorandom seed. The shuffling system is also inherently deterministic and would not change in consecutive executions if the index and pose weights maintain their previous values. Changing just the index value will vary the produced sequence of poses. This system works well because the pseudorandom sequences used in this version of Simon are much shorter than in the original game. Larger sequences could end in possible noticeable loops.

In the shuffling system, the idea is to find the pose to be executed by adding the index and the weight of the last executed pose to the last identification number (ID) of the pose. To visualize this, Table 1 shows an example of poses with ID and weights:

| pose_id | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| weight | 2 | 0 | 1 | 2 | 3 | 1 |

Table 1. Different poses with identification numbers and their associated weight.

Assuming that the index is 1: if the last time we chose the pose with ID 2, the next pose will be found by adding weight+index to the ID number (2+1+1). So the next pose to be executed is 4. However, if that combination exceeds the defined number of poses in the PDDL problem, this combination value is reduced according to the maximum number of poses. So, in Table 1, when the last executed pose is 4, the next pose ID is 8 and then decremented to 3 (equivalent to 8 mod 5). This process is visualized in Figure 4 with the corresponding actions planned by the automated planner: execute-pose, findnext-id and nextpose-id.

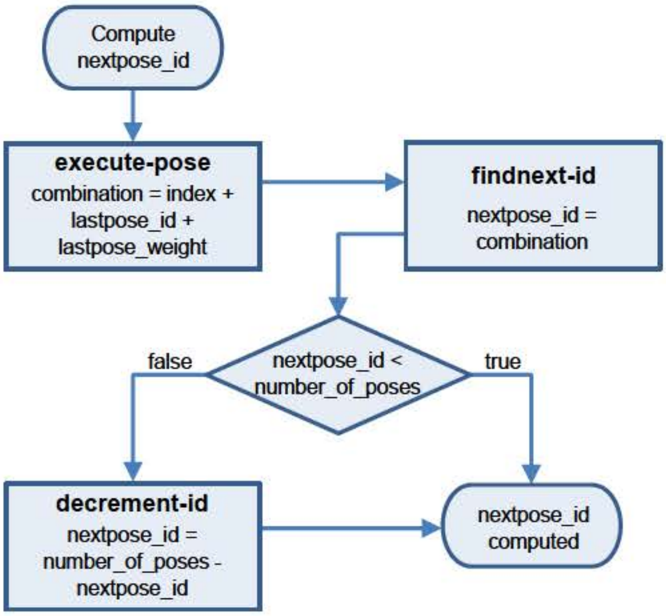



Fig. 4. Flowchart summarizing the process of computing the new pose identifier.

### 5.2. *Simon in the NAOTherapist architecture*

The NAOTherapist architecture is based on RoboComp components,[30] which are individual processes that can be programmed in different languages and even executed in different physical machines. The planning of the game session, its execution and monitoring follow the strategy of the PELEA subsystem as shown in Figure 5.[32] The Executive component receives the perception data and send them to PELEA to obtain the next planned action. Then, Executive translates this action into instructions to be executed by the robot, if it is required.

If the previous plan is no longer valid due to unexpected events, like the user leaving the playing area in front of the Kinect sensor, PELEA can detect it using its Monitoring module. Then it uses the Decision Support module to obtain a new plan to reach the goals of the PDDL problem (finish the game in this case). Decision Support uses the Metric-FF automated planner.[35] Execution module serves as an interface between PELEA and the other components of the architecture. The actions managed by the Simon domain state as follows:

- detect-patient: if patient is not detected, try to detect patient.

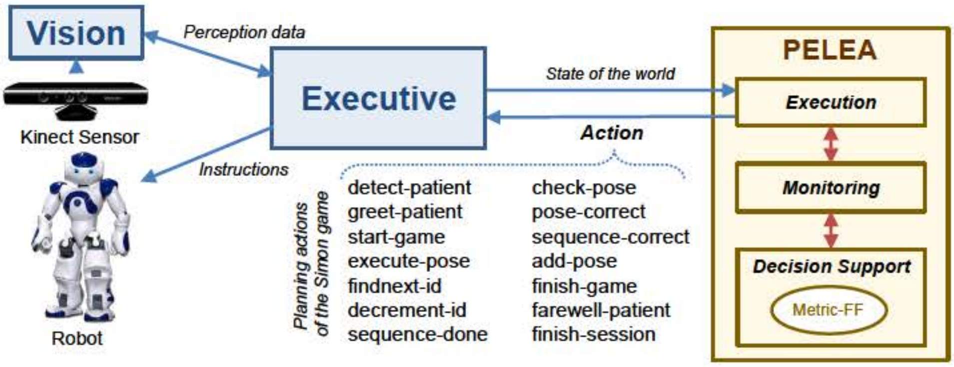



Fig. 5. List of planning actions of the Simon game domain provided by PELEA and translation of these actions into instructions by the Executive component to the robot while preserving the state of the world.

- greet-patient: if a patient is detected but not yet greeted, greet the patient.
- start-game: if the patient is greeted, start the game by initializing the corresponding numeric fluents.
- execute-pose: as long as the number of poses is less than or equal to the turn counter, execute the pose whose ID was decided by findnext-id.
- findnext-id: after a pose was executed and while it is not the end of the sequence, assign a new value as the next pose to be executed.
- decrement-id: if the next pose ID is higher than the number of poses defined in the PDDL problem, decrement this number by subtracting the number of poses.
- sequence-done: when the pose counter is equal to the number of poses to be done on that particular sequence, notify the program that this round is done and finish the robot's turn.
- check-pose: check if the pose is correct by comparing the arm postures of the patient with the pose configurations.
- pose-correct: if the pose is correct, increment the pose counter and update the necessary numeric fluents.
- sequence-correct: when all poses are checked and all of them are correct, start the next turn of the robot.
- add-pose: increment the turn counter so that one pose more can be executed on the next sequence. Initialize and update the numeric fluents for the new round.
- finish-game: when the number of pre-determined turns to be played is done, finish the game.
- farewell-patient: after the game is finished, farewell the patient to notify him that the game is over.
- finish-session: finish the session after the robot said goodbye to the patient.

The pose recognition mechanism relies in a component of the NAOTherapist architecture that has been completely reused, including the admissible angle thresholds.[12] Each pose is accepted when the system detects that it has been maintained for a certain amount of time. The perceived pose is compared against the expected one, which is stored in an internal knowledge base. The Simon game mechanics explained in this manuscript replaced the original mirror game developed for NAOTherapist.

## 6. Evaluation of the Human-Robot Interaction

This section details the evaluation of the developed system in terms of Human-Robot Interaction. For the evaluation, statistical data were collected from log files of every execution and user feedback was gathered through questionnaires completed after the test sessions.

### 6.1. *Experimental setup*

Two different evaluations were performed: one with adults and another with schoolchildren. The experiments were carried out in a room with the robot on the floor and the Kinect camera elevated behind it. The user was 3 meters in front of the robot. After the session, the users filled a questionnaire in a nearby table. The computer in which all the developed code was executed is an Intel Core2 Duo at 1.86 GHz with 2 GB of RAM memory. This computer is fast enough to maintain the planning time always under 2 seconds, although times near that threshold were only extreme cases that could be caused by other operations of the computer. This ensures that the possible execution delays are due to the design of the interaction, instead of the platform.

The first study was done with 14 adults varying in age from 20 to 50, all playing the game for the first time. After this initial evaluation to tune the game flow, the system was better prepared for testing with children. Then the game was played in a school along two different sessions with 56 healthy children of 5 years old. The first evaluation used the whole pose catalogue of NAOTherapist,[12] but it was too complicated for children of such age, so we reduced it to the subset shown in Figure 6.

The tests presented in this manuscript evaluate three aspects of the system: the quality of the human-robot interaction, the robot performance in the game and the game format itself. The game format is evaluated mostly in terms of difficulty of the poses and duration of the session, whereas the robot performance is analyzed to check whether the provided feedback from the robot is enough and if it is both suitable and easy to comprehend for users. For the human-robot interaction part, we questioned the approach of the robot and the feelings to get the user's perception.

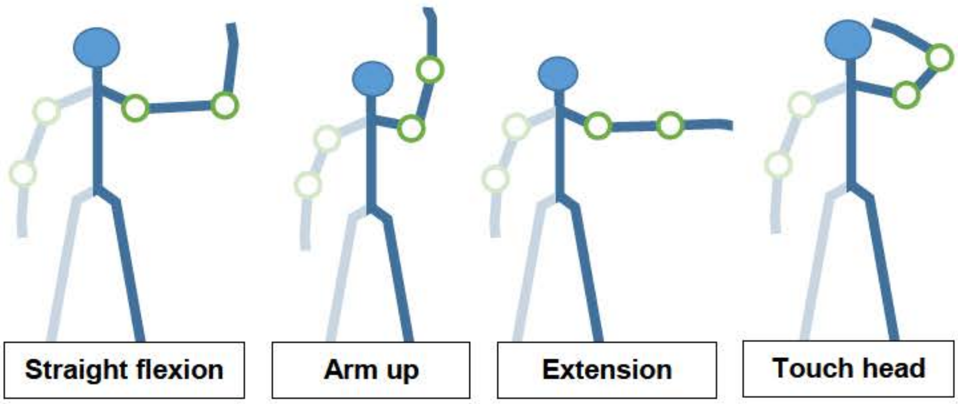



Fig. 6. Types of arm postures considered in the evaluation with children.

### 6.2. *First evaluation: adults*

In this initial study, users played two games with the robot, both containing five turns of poses. Table 2 displays the statistics of participants. This evaluation used the full pose catalogue of NAOTherapist, allowing only poses with one arm or symmetrical ones. With 7 different types, there are 21 possible arm combinations that could appear in a sequence. It can be observed that the arithmetic mean of the second game duration is less than the first one. The reminders average seems to be moderately high, but it should be noted that it was mostly caused by some few users, thus the also high standard deviation.

| Participants | Age | Maximum turns (poses) | Game-1 duration (s) | Game-2 duration (s) | Reminders | Skipped poses | Fails |
|---|---|---|---|---|---|---|---|
| 14 | 32±11 | 2x5 (2x15) | 235.2±32 | 231.8±34 | 4.23±2.9 | 1.31±1.3 | 7.69% |

Table 2. Average results from the first evaluation with adults.

According to data acquired from the statistics of the test sessions, experience and difficulty are the balancing features of the game. All participants were playing the game for the first time, so their experience level is known to be zero at the beginning of the test process, then raised with the testing of the first and the second game during the sessions. Thus it was expected from the majority to perform better in the second game since they gained familiarity about the game and its rules. The second game had more difficult poses, but the difference between the average durations of two games was only 3.4 seconds, second game being shorter. This gives the clue that when the two games are presented for the first time and are on the same difficulty level, second game will be far easier for the player, concluding in a shorter duration. In the questionnaire answers, people shared that they found the first game easier to accomplish but admitted to have enjoyed the second one better, presumably engaged with their higher level of familiarity with the game.

Despite the similarity of the durations, in the second game participants clearly performed better in terms of achieving the poses and following the instructions of the robot. Less reminders were needed and the number of pose skips fell to zero. However, it should be considered that the number of skips were only 17 of over 400 executed poses throughout the whole test session. Also the statistics of pose reminders indicates that half the participants never skipped a pose. Furthermore, of the 56 poses reminded to the participants, 70% of the time they were able to correct it before skipping. This is good because it indicates that users reasoned about their posture.

In the questionnaires, the majority of the answers obtained on human-robot interaction quality is compatible with the project goals, although the users had different opinions about the difficulty of the game and the speed of the execution. The idea about whether this particular game could be played by children was positive for all participants. They also pointed out its good communication and easy-going manner as reasons to be used with children, while the rest referred to the appearance or other physical aspects.

### 6.3. *Second evaluation: children*

On the second study, the game was tested with children of 5 years old. After taking into account the tests with adults, the difficulty was reduced by using the posture subset of Figure 6 and making both arms to have the same posture. This way, all poses were symmetrical. This reduced the pose combinations to only 4, like the original Simon game. Some more instructions were added at the beginning to explain the game mechanics better. The tests started with an increased number of maximum poses of 7. However, after observing that many children had a hard time [illegible] of 5. Table 3 displays the statistics of the test sessions. The fail rate at finishing the game decreased drastically when the number of turns lowered to 5 and also the skipped poses, that are further from the limit of 3. So this indicates that the number of poses play an important role in the feasibility of the game and the number must be determined carefully.

| Participants | Age | Maximum turns (poses) | Game duration (s) | Reminders | Skipped poses | Fails |
|---|---|---|---|---|---|---|
| 24 | 5±0 | 7 (28) | 434.0±85 | 22.83±7.4 | 2.79±1.6 | 29.17% |
| 32 | 5±0 | 5 (15) | 274.0±33 | 10.81±4.5 | 1.00±1.1 | 18.75% |

Table 3. Average results from the second evaluation with children.

The questionnaire used in this evaluation is derived from the one developed by Heerink et al.[36] Questions are answered using a 5-point Likert scale except some open questions (Q8, Q9, Q10 and Q14). Here, we have normalized the values and, in some cases, reversed them so that 1 is always the most desirable outcome and 0 the least. The questions and their identifiers can be found in Appendix A.

Figure 7 shows the user feedback, where each bar represents the normalized arithmetic mean of the Likert questions and lines indicate the standard deviation. Human-robot interaction part of the game mostly met the minimum expected criterion (higher than 0.5) and even had a high rank close to 1. Even though there are some questions that could not achieve the 0.75, they were not in the undesirable region. This tells us that, in the overall picture, the children enjoyed the experience of playing with the robot.

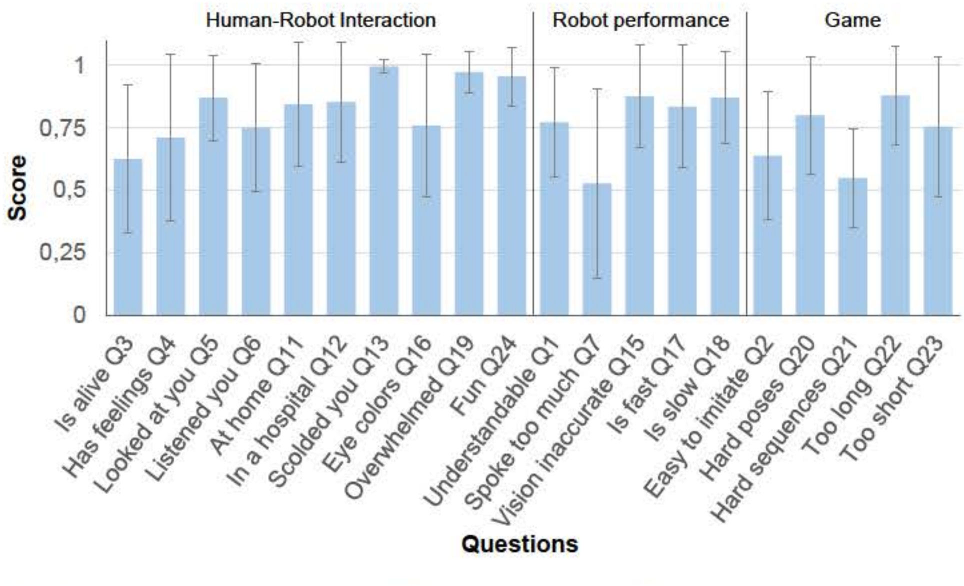



Fig. 7. Questionnaire outcomes with arithmetic means and standard deviation. Q8, Q9, Q10 and Q14 not included because they are open-ended questions.

One open question (Q14) was about what else they like to play with the robot. Their answers include hide and seek, tag or even soccer: mostly games that require running or at least to perform active movements. These very interactive games shows that children saw the robot as a play companion that is capable of doing more than just talking, standing and showing poses. This impression is believed to be built with the help of the audio and motion feedback about the movements of the child during the game. Also, we asked for names for the robot (Q9). They sometimes chose their friend's name or the name of a cartoon character which probably they are fond of.

One part of the questionnaire (Q8) asked for how would they describe the robot itself. A number of adjectives was provided and children had to choose among them. Half of the adjectives represented good values. At the same time, half of them were social adjectives that could be applied to a user more than to a machine.

As seen in Figure 8, the most frequently chosen word is happy, followed by loving, clever and friendly. This is for sure the desired outcome: that the robot had a favorable impact on the children. Other than the [illegible] adjectives, there are occurrences of some adjective that can be considered [illegible]ve, like difficult, impatient and angry. Even though the total of their occur[illegible]rly low, they are [illegible] word angry, the same child also chose loving and happy for the robot, so it may be either a misuse of the word or caused by the stress of remembering poses that seem difficult. Difficult and impatient words are believed to address the game format.

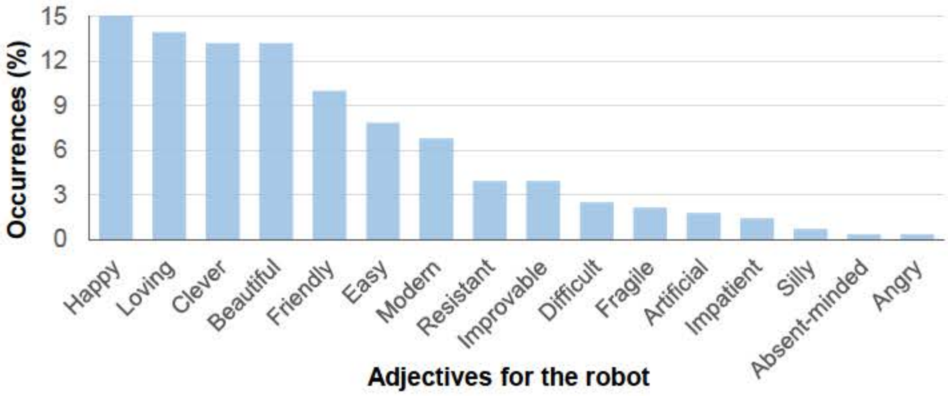



Fig. 8. Percentage of adjectives chosen by children.

Performance-wise, nearly all of participants thought the speed of the robot to execute the actions is just right, but some participants found it too fast. This and the occurrences of the aforementioned adjectives are a sign to further decrease the difficulty of the poses at least for users that are having a hard time catching up or remembering them. This problem could be addressed by collecting data about the user and the session progress, and customizing the difficulty of the new pose as well as the waiting time between poses accordingly.

Coming to evaluate the game format, despite the increased number of poses at the beginning of the evaluations, the children answered that the duration of the game was neither too long or too short. This outcome validates the attained data from the initial evaluation that 5 turns is enough but can be short for children. This is a good reference point to determine the number of sequences to be played in a real training session.

In the overall picture, the weakest points of this game while playing with the children are discovered to be the difficulty of the poses. Children found some particular poses hard to remember or perform and stated it as their least favorite thing about the game.

## 7. Discussion of the HRI results

Some of the points to be noted in the execution of the game and the interaction are the false perception of self pose of humans (proprioception), and also player's emotions during the interaction.

Regarding the first point, it was observed that when the poses expected from the user are more difficult than an intermediate level, then a problem occurs that was not taken into account before: the performance difference of the robot and the human. It was noticed, especially during the tests, that people experienced difficulty comprehending the posture they have to maintain. Sometimes they believed that it was a mistake of the robot while perceiving their pose. This is a consequence of the faulty human self spacial perception. For instance, raising the arm of a shape of L (straight flexion in Figure 6) may seem easy to the player but mostly it is failed due to not enough lifting of the elbow.

As a solution, a mirrored correction can be implemented as it is in NAOTherapist. In this wise, robot will imitate the actual players' pose and show the right pose consecutively. This way they will be noticed about their true posture, thus knowing where they were mistaken and what they should do to correct it. Correction of that kind would not be suitable for someone who merely forgot the pose. It should be used only when the player is close to achieve the pose but he cannot do it in the given time.

The second point is about the anxiety, frustration, boredom and laziness feelings reported by some participants after the session. All participants in these evaluations faced the robot for the first time, so many of these feelings could be solved just by playing more and adapting to the mechanics of the game. Even though their average is fairly low, they should be taken into account and worked on, for the benefit of the interaction. Frustration, boredom and laziness feelings are admitted to be caused mostly by the pose correction process. Failing the pose several times could be irritating and may lead the user to lose interest in the game. Those feelings could be overcome by implementing the mirrored correction explained before. Anxiety is believed to be caused by playing a recently implemented format of a game with a robot that people have not communicated with before. Furthermore, the session could start with shorter and easier warm-up games to help the players acquire experience and feel more confident around the robot. Anxiety probably could be alleviated by a brief introductory session too. The red color of the eyes could be a factor of stress in case of failures, although the questionnaire outcomes do not indicate that. Anyway, is therefore worthwhile to do a future evaluation with eye colors ranging from white to green, although the color gradient is smaller in the latter case.

Thinking quality wise for a long-term engagement of the children, the reversed version of the game might also be considered in the near future. Although it has not been implemented in the project, the idea is to have the user performing poses and the robot imitating them in incremental sequences. This will surely improve the experience of communicating with the robot for the user, leading to a more cooperative and natural interaction.

All this was achieved within the original NAOTherapist architecture, demonstrating its domain generalization capabilities and the advantages of reusing a previous architecture instead of rewriting ad hoc solutions from scratch. Moreover, the developed declarative PDDL domain for Simon can be easily extended to address many more actions without maintaining a complex state machine.

## 8. Conclusions

This work presents a novel approach of the electronic Simon game which replaces colors by upper-limb poses while using a social robot NAO to drive the game session. Focusing on the human-robot interaction aspect, we explain the Simon game development to further learn from the interactive game format on how to improve its human-robot interaction.

The robot-based Simon game is modelled using Automated Planning techniques into the NAOTherapist architecture. This work demonstrates its capability to be extended to other domains and its utility for further developments. The PELEA automated planning subsystem is in charge of planning, monitoring and executing the session as used in the original architecture. Also, although it was thought to be an assistive robotic application for people with disabilities, it can be carried on to be an application for a broader audience.

Based on the evaluations from the test participants, we have demonstrated that the interaction with the new developed game worked very well with children and that they enjoyed the time they spent with the robot. According to the outcome of these evaluations, the Simon game is a promising prototype to check its therapeutic benefits in a clinical environment.

## Appendix A. Children's Questionnaire

Q1. Was it easy to understand how to play with the robot?
Q2. Was it easy to imitate the poses shown by the robot?
Q3. Do you think the robot is alive?
Q4. Do you think the robot has feelings?
Q5. Did you have the impression that the robot looked at you?
Q6. Did you have the impression that the robot listened to you?
Q7. Do you think the robot spoke too much?
Q8. Choose 5 adjectives to describe the robot.
Q9. How would you like to name the robot?

Q10. How old do you think the robot is?
Q11. Would you like to have this robot at home?
Q12. Would you like to be treated by this robot in a hospital?
Q13. Do you think that the robot scolded you?
Q14. Which other games would you want to play with the robot?
Q15. Did the robot make you repeat a correct pose?
Q16. Were the colors of the eyes useful for doing the exercises?
Q17. Do you think that the pace of the game was too fast?
Q18. Do you think that the pace of the game was too slow?
Q19. Did you feel overwhelmed when the robot played with you?
Q20. Were the poses too difficult to imitate?
Q21. Were the poses too difficult to remember?
Q22. Do you think that the time playing with the robot was too long?
Q23. Do you think that the time playing with the robot was too short?
Q24. Did you have fun playing with the robot?
Q25. What did you like the most while playing Simon with the robot?
Q26. What did you like the least while playing Simon with the robot?

## Acknowledgements

This work is partially funded by grant TIN2012-38079-C03-02 and TIN2015-65686-C5-1-R of Spanish Ministerio de Economía y Competitividad. We also want to thank the Joan Miró school of Leganés for their assistance with the evaluations, to the teachers and the management team for their support, and specially to all the children who kindly participated in the evaluation and enjoyed playing with our robots.

## References

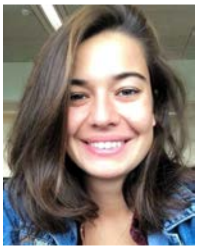
1. K. R. Ginsburg, "The Importance of Play in Promoting Healthy Child Development and Maintaining Strong Parent-Child Bonds," *Pediatrics*, vol. 119, pp. 182–191, Jan. 2007.
2. T. M. Fleming, L. Bavin, K. Stasiak, E. Hermansson-Webb, S. N. Merry, C. Cheek, M. Lucassen, H. M. Lau, B. Pollmuller, and S. Hetrick, "Serious Games and Gamification for Mental Health: Current Status and Promising Directions," *Frontiers in Psychiatry*, vol. 7, p. 215, Jan. 2017.
3. J. Janssen, O. Verschuren, W. J. Renger, J. Ermers, M. Ketelaar, and R. van Ee, "Gamification in Physical Therapy: More Than Using Games," *Pediatric Physical Therapy*, vol. 29, pp. 95–99, Jan. 2017.
4. D. Feil-Seifer and M. J. Matarić, "Defining Socially Assistive Robotics," in *Proceedings of the 9th International Conference on Rehabilitation Robotics (ICORR)*, pp. 465–468, IEEE, June 2005.
5. A. Pereira, C. Martinho, I. Leite, and A. Paiva, "iCat, the chess player: the influence of embodiment in the enjoyment of a game," in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, (Estoril, Portugal), pp. 1253–1256, May 2008.

6. E. S. Kim, L. D. Berkovits, E. P. Bernier, D. Leyzberg, F. Shic, R. Paul, and B. Scassellati, "Social Robots as Embedded Reinforcers of Social Behavior in Children with Autism," *Journal of Autism and Developmental Disorders (JADD)*, vol. 43, pp. 1038–1049, May 2013.
7. P. Pennisi, A. Tonacci, G. Tartarisco, L. Billeci, L. Ruta, S. Gangemi, and G. Pioggia, "Autism and Social Robotics: A Systematic Review," *Autism Research*, vol. 9, pp. 165–183, Feb. 2016.
8. K. Dautenhahn and I. Werry, "Towards interactive robots in autism therapy: Background, motivation and challenges," *Pragmatics & Cognition*, vol. 12, pp. 1–35, June 2004.
9. W. A. Bainbridge, J. Hart, E. Kim, and B. Scassellati, "The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents," *International Journal of Social Robotics*, vol. 3, pp. 41–52, Oct. 2011.
10. B. Robins, E. Ferrari, K. Dautenhahn, G. Kronrief, B. Prazak-Aram, G. J. Gelderblom, F. Caprino, and E. Laudana, "Developing Scenarios for Robot Assisted Play Informed by User Panels and Field Trials," *International Journal of Human-Computer Studies (IJHCS)*, vol. 68, pp. 873–898, Jan. 2010.
11. J. C. González, J. C. Pulido, and F. Fernández, "A three-layer planning architecture for the autonomous control of rehabilitation therapies based on social robots," *Cognitive Systems Research (CSR)*, vol. 43, pp. 232–249, June 2017.
12. J. C. Pulido, J. C. González, C. Suárez-Mejías, A. Bandera, P. Bustos, and F. Fernández, "Evaluating the Child–Robot Interaction of the NAOTherapist Platform in Pediatric Rehabilitation," *International Journal of Social Robotics*, vol. 9, pp. 343–358, June 2017.
13. M. Turp, J. C. Pulido, J. C. González, and F. Fernández, "Playing with Robots: An Interactive Simon Game," in *Proceedings of the 16th Conference of the Spanish Association for Artificial Intelligence (CAEPIA), RSIM workshop*, (Albacete, Spain), pp. 1085–1095, Nov. 2015.
14. M. Ghallab, D. Nau, and P. Traverso, *Automated Planning: Theory & Practice*. Elsevier, May 2004.
15. M. Fox and D. Long, "PDDL2.1: An Extension to PDDL for Expressing Temporal Planning Domains," *Journal of Artificial Intelligence Research (JAIR)*, vol. 20, pp. 61–124, Dec. 2003.
16. A. Papangelis, G. Mouchakis, A. B. Texas, D. Kosmopoulos, V. Karkaletsis, and F. Makedon, "A Game System for Remote Rehabilitation of Cerebral Palsy Patients," in *Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA)*, pp. 19:1–19:9, ACM, June 2012.
17. C. McMurrough, S. Ferdous, A. Papangelis, A. Boisselle, and F. M. Heracleia, "A Survey of Assistive Devices for Cerebral Palsy Patients," in *Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA)*, pp. 17:1–17:8, ACM, June 2012.
18. H. Kose, N. Akalin, and P. Uluer, "Socially Interactive Robotic Platforms as Sign Language Tutors," *International Journal of Humanoid Robotics (IJHR)*, vol. 11, p. 1450003, Mar. 2014.
19. H. Lehmann, I. Iacono, B. Robins, P. Marti, and K. Dautenhahn, "'Make It Move': Playing Cause and Effect Games with a Robot Companion for Children with Cognitive Disabilities," in *Proceedings of the 29th Annual European Conference on Cognitive Ergonomics (ECCE)*, pp. 105–112, ACM, Aug. 2011.
20. A. G. Brooks, M. Berlin, and J. Gray, "Untethered Robotic Play for Repetitive Physical Tasks," in *Proceedings of the 2nd International Conference on Advances in Com-*

*puter Entertainment Technology (ACE)*, pp. 27–34, ACM, June 2005.
21. L. Roberts, H. W. Park, and A. M. Howard, "Robots and Therapeutic Play: Evaluation of a Wireless Interface Device for Interaction with a Robot Playmate," in *Proceedings of the 34th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS)*, pp. 6475–6478, IEEE, Aug. 2012.
22. D. Nau, T.-C. Au, O. Ilghami, U. Kuter, J. W. Murdock, D. Wu, and F. Yaman, "SHOP2: An HTN Planning System," *Journal of Artificial Intelligence Research (JAIR)*, vol. 20, pp. 379–404, Dec. 2003.
23. C. Galindo, J. González, and J. A. Fernández-Madrigal, "An Architecture for Cognitive Human-Robot Integration. Application to Rehabilitation Robotics," in *Proceedings of the 2nd International Conference on Mechatronics and Automation*, vol. 1, pp. 329–334, IEEE, July 2005.
24. M. La Placa, H. Pigot, and F. Kabanza, "Assistive planning for people with cognitive impairments," in *Proceedings of the Workshop on Intelligent Systems for Assisted Cognition, IJCAI conference*, (Pasadena, California, USA), pp. 69–76, July 2009.
25. J. Fasola and M. Matarić, "Robot Exercise Instructor: A Socially Assistive Robot System to Monitor and Encourage Physical Exercise for the Elderly," in *Proceedings of the 19th International Symposium in Robot and Human Interactive Communication (ROMAN)*, (Viareggio, Italy), pp. 416–421, Sep. 2010.
26. D. Feil-Seifer and M. Matarić, "A Simon-Says Robot Providing Autonomous Imitation Feedback Using Graded Cueing," in *Proceedings of the 12th International Meeting for Autism Research (IMFAR)*, (Toronto, Ontario, Canada), May 2012.
27. A. M. Howard, "Robots Learn to Play: Robots Emerging Role in Pediatric Therapy," in *Proceedings of the 26th International Florida Artificial Intelligence Research Society Conference (FLAIRS)*, (St. Pete Beach, USA), May 2013.
28. V. L. Calderita, J. L. Manso, P. Bustos, C. Suárez-Mejías, F. Fernández, and A. Bandera, "THERAPIST: Towards an Autonomous Socially Interactive Robot for Motor and Neurorehabilitation Therapies for Children," *JMIR Rehabilitation and Assistive Technologies (JRAT)*, vol. 1, p. e1, Oct. 2014.
29. R. Sun, *Duality of the Mind - A Bottom-up Approach Toward Cognition*. Lawrence Erlbaum, 2001.
30. L. Manso, P. Bachiller, P. Bustos, P. Núñez, R. Cintas, and L. Calderita, "RoboComp: A Tool-Based Robotics Framework," in *Simulation, Modeling, and Programming for Autonomous Robots*, vol. 6472 of *Lecture Notes in Computer Science*, pp. 251–262, Springer Berlin Heidelberg, 2010.
31. M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: an open-source Robot Operating System," in *Proceedings of the Workshop on Open Source Software in Robotics, ICRA conference*, (Kobe, Japan), May 2009.
32. V. Alcázar, C. Guzmán, D. Prior, D. Borrajo, L. Castillo, and E. Onaindia, "PELEA: Planning, Learning and Execution Architecture," in *Proceedings of the 28th Workshop of the UK Planning and Scheduling Special Interest Group (PlanSIG)*, (Brescia, Italy), Dec. 2010.
33. M. Chekaf, N. Gauvrit, A. Guida, and F. Mathy, "Chunking in Working Memory and its Relationship to Intelligence," in *Proceedings of the 37th Annual Meeting of the Cognitive Science Society (CogSci)*, July 2015.
34. M. H. Gendle and M. R. Ransom, "Use of the Electronic Game SIMON as a Measure of Working Memory Span in College Age Adults," *Journal of Behavioral and Neuroscience Research*, vol. 4, pp. 1–7, 2006.
35. J. Hoffmann, "The Metric-FF Planning System: Translating "Ignoring Delete Lists" to

Numeric State Variables," *Journal of Artificial Intelligence Research (JAIR)*, vol. 20, pp. 291–341, Dec. 2003.

36. M. Heerink, M. Díaz, J. Albo-Canals, C. Angulo, A. Barco, J. Casacuberta, and C. Garriga, "A field study with primary school children on perception of social presence and interactive behavior with a pet robot," in *Proceedings of the 21st International Symposium on Robot and Human Interactive Communication (ROMAN)*, pp. 1045–1050, IEEE, Sep. 2012.



**Mısra Turp** studied Computer Science at Sabancı Üniversitesi (SU), Turkey. In 2015 she was a visitor student at Universidad Carlos III de Madrid (UC3M), Spain, working on Automated Planning and plan execution for Autonomous Robotics within the NAOTherapist project. She continued working on Humanoid Robotics for her master's thesis and, in 2019, she received her M.S. degree focused on Artificial Intelligence from Universiteit van Amsterdam (UvA), The Netherlands. She also explored the multi-agent reasoning field with a scientific publication in the 2017 IEEE International Conference on Robotics and Automation (ICRA). Currently, she is working as a data scientist in Amsterdam for IBM Netherlands.

**José Carlos González** received his M.S. degree in 2014 from the Universidad Carlos III de Madrid (UC3M), Spain. Currently, he is finishing there his Ph.D. studies of Computer Science and Engineering. He is also one of the main developers of NAOTherapist. This project about Artificial Intelligence and Socially Assistive Robotics uses a cognitive architecture which is part of the outcome of his doctoral thesis. In 2012 he was a visitor student at Karlsruher Institut für Technologie (KIT), Germany, and in 2017 he was a visitor scholar in the Machine Learning Department at Carnegie Mellon University (CMU), USA. His current research interests include control architectures for Autonomous Robotics, Automated Planning and Machine Learning techniques.

**José Carlos Pulido** is a last year Ph.D. student in the Computer Science Department at Universidad Carlos III de Madrid (UC3M). He received there his B.D. in Computer Science and Engineering. In 2012 he was studying at Karlsruher Institut für Technologie (KIT) in Germany, where he finished his bachelor's thesis focused on Industrial Robotics and Computer Vision. In 2014 he continued his studies at UC3M receiving his M.D. in Computer Science and Technology. In 2018 he was a visiting scholar at the Interaction Lab of the University of Southern California (USC). His line of research is focused on the design and development of Motor Rehabilitation Therapies based on Socially Assistive Robotics.

**Fernando Fernández** is a faculty of the Computer Science Department of Universidad Carlos III de Madrid (UC3M) since 2005. He received his Ph.D. degree in Computer Science from UC3M in 2003. In 2000, Fernando was a visiting scholar at the Center for Engineering Science Advanced Research at Oak Ridge National Laboratory, Tennessee. He was also a postdoctoral fellow at the Computer Science Department of Carnegie Mellon University (CMU) since october 2004 until december 2005. He is the recipient of a pre-doctoral FPU fellowship award from Spanish Ministry of Education (MEC), a Doctoral Prize from UC3M, and a MEC-Fulbright postdoctoral Fellowship. He has more than 60 journal and conference papers, mainly in the fields of Machine Learning and Automated Planning. He is interested in decision making tools and intelligent systems that operate in continuous and stochastic domains, such as Socially Assistive Robotics.